White Paper

IMPROVING VMWARE DISASTER RECOVERY WITH EMC RECOVERPOINT

Applied Technology

Abstract

EMC® RecoverPoint provides full support for data replication and disaster recovery for VMware® ESX® Server and ESX servers' virtual machine clients. This white paper describes how RecoverPoint can be utilized to provide local and remote data protection and recovery for VMware ESX environments. It also covers the supported configurations available for VMware ESX Server and ESX virtual machine environments and the integration of RecoverPoint with VMware vCenter™ Site Recovery Manager.

December 2011

Part Number h2352.6

Table of Contents

Executive summary

EMC® RecoverPoint is an advanced enterprise-class disaster recovery solution supporting heterogeneous storage and server environments. RecoverPoint provides bi-directional local and remote data replication across any distance, and utilizes continuous data protection technologies to provide consistent point-in-time recovery. RecoverPoint helps customers accelerate protection and provides operational and disaster recovery of their VMware® Infrastructure, without impacting production environments. RecoverPoint is ideally suited for replicating and protecting physical and virtual server environments. RecoverPoint 3.4 and later supports replicating datastores managed by vSphere 5. The latest RecoverPoint Storage Replication Adapter (2.0 or later) is required for VMware vCenter Site Recovery Manager 5.0.

Introduction

Server virtualization technology allows one physical server platform to run multiple virtual machines simultaneously. Many customers have taken advantage of server virtualization, such as that provided by VMware ESX®, to consolidate their server infrastructure and simplify their disaster recovery platforms. These customers may have also invested in enterprise-class SAN switches, such as the EMC Connectrix® ED-DCX-4S-B and MDS 9000 director family, along with a SAN-based storage infrastructure, which support their primary data center and disaster recovery sites. This leads to some challenges when it comes to managing data protection for their local and remote data centers, especially for applications running on a virtual machine in an ESX server.

This white paper describes how customers can utilize RecoverPoint to enhance the disaster recovery and data protection capabilities of their physical and virtualized VMware servers with local and remote replication.

Audience

This white paper is targeted to VMware storage and server administrators, IT managers, and storage professionals, as well as integrators, consultants, and distributors.

EMC RecoverPoint

RecoverPoint is an out-of-band appliance-based product, designed with the performance, reliability, and supportability required for enterprise applications. Running on a cluster of tightly coupled servers, RecoverPoint's high-availability design ensures that the failure of a single appliance will not affect the data protection of VMware ESX Server. RecoverPoint utilizes write splitters that reside on the VNX™ series, CLARiiON® CX3, and CX4 arrays, in the fabric or on the host, to monitor and intercept writes to protected volumes so that a copy of the write can be sent to the RecoverPoint appliance for further processing.

RecoverPoint provides a full-featured replication and continuous data protection solution for VMware ESX servers. For remote replication, it utilizes synchronous replication with a zero recovery point objective (RPO), or asynchronous replication with a small RPO. This enables RecoverPoint to protect the VMware ESX platform from data corruption and guarantees recoverability with little to no data loss. For local protection, it preserves every write, allowing data recovery to any point in time.

RecoverPoint CDP, CRR, and CLR

EMC RecoverPoint provides local and remote replication. Local replication of a LUN is provided through the EMC RecoverPoint continuous data protection (CDP) capability and remote replication of a LUN is provided though the RecoverPoint continuous remote replication (CRR) capability. Additionally the same LUNs can be protected locally and remotely. This capability is called continuous local and remote (CLR) data protection.

The innovative technology of RecoverPoint supports flexible levels of protection without distance limitations or performance degradation. RecoverPoint's technology offers a fine-grain recovery of application data and reduces the recovery point to zero. Users can recover their data to any point in time, eliminating the need to invest in physical recovery of data damaged because of server outages, data corruption, software errors, viruses, or common user errors.

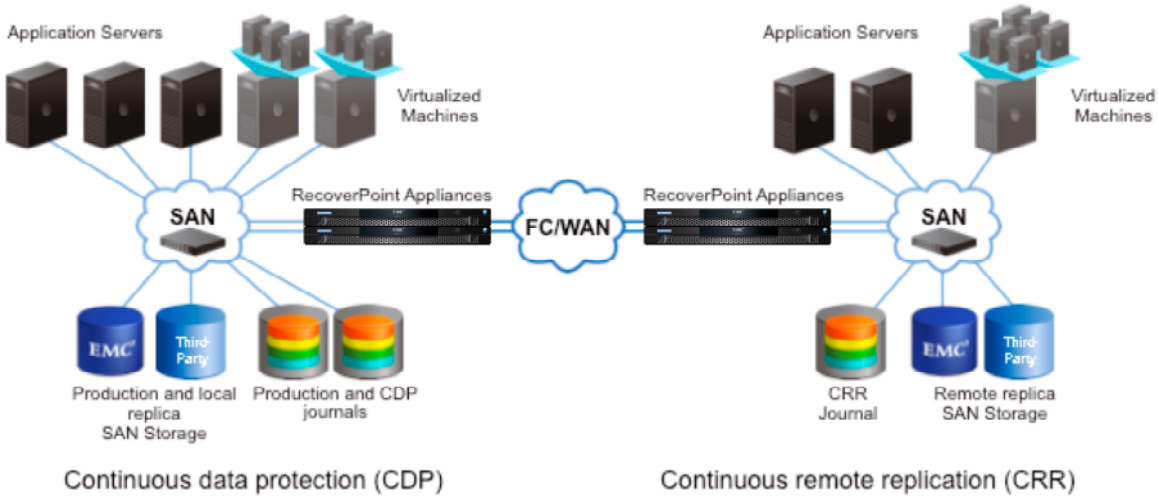

Figure 1. EMC RecoverPoint architecture overview

All of these capabilities help the customer achieve a dramatically lower total cost of ownership (TCO) compared to other host- or array-based replication solutions. Recovery testing is also made easier because of the ability to access the replicated data at the local or remote site for recovery or integrity testing purposes without interrupting the replication or the ongoing data center operations.

Write-splitting technologies

Array-based write splitting

RecoverPoint supports array-based write splitting on Symmetrix VMAXe, VNX series and CLARiiON CX3 and CX4 arrays. The array-based write splitter is included with the Symmetrix VMAXe and the VNX while the CLARiiON write splitter is a feature of CLARiiON FLARE® 26, 28, 29, and 30. The array-based write splitter is supported on the Symmetrix VMAXe series, VNX series and CLARiiON CX3 and CX4 arrays as well as for block volumes available on Celerra® unified arrays, including the NS20, NS40, NS80, NS-120, NS-480, and NS-960 models. The write-splitting function operates in each storage processor where it monitors writes to protected volumes and ensures that the RecoverPoint appliance receives a copy of the write. Up to four RecoverPoint clusters can be attached to the same VMAXe series, VNX series or CLARiiON array. A benefit of this is that multiple production sites can use the same array such as might exist in a shared disaster recovery site or at a service provider's site.

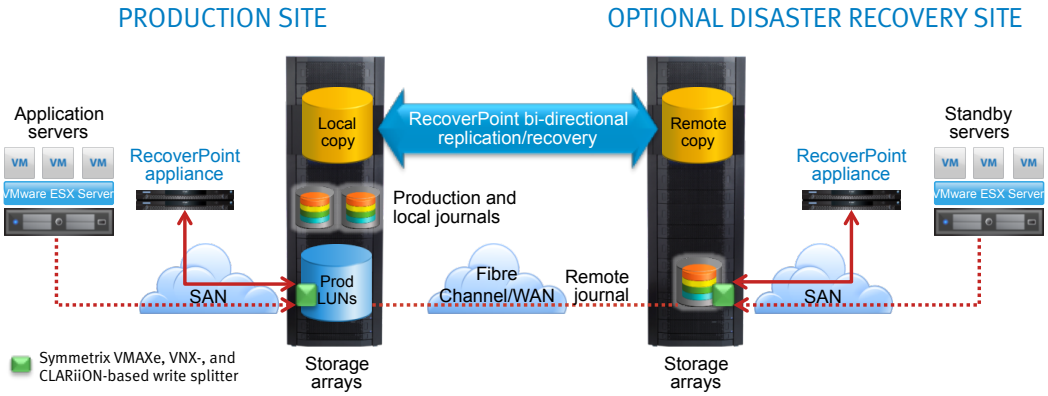

Figure 2. Array splitter architecture

For VMware virtual machines, this enables replication of VMFS 3 and VMFS 5 and physical RDM/P (RDM/P) volumes without the cost or complexity of an intelligent fabric implementation. The splitter supports Fibre Channel (FC) volumes for all arrays as well as iSCSI volumes presented by the VNX or CLARiiON array to any host, including to an ESX server.

Host-based write splitting

For VMware, RecoverPoint provides a host-based write splitter (also called a KDriver) for Windows Server platforms. The KDriver is installed on each Windows virtual machine where it operates above any multipath driver, but below the file system and volume management layers. The KDriver monitors writes and ensures that a copy of all writes to a protected volume is sent to the RecoverPoint appliance. Since the KDriver runs in the virtual machine, the only volumes that can be replicated by RecoverPoint would be SAN volumes attached to the virtual machine in RDM/P mode.

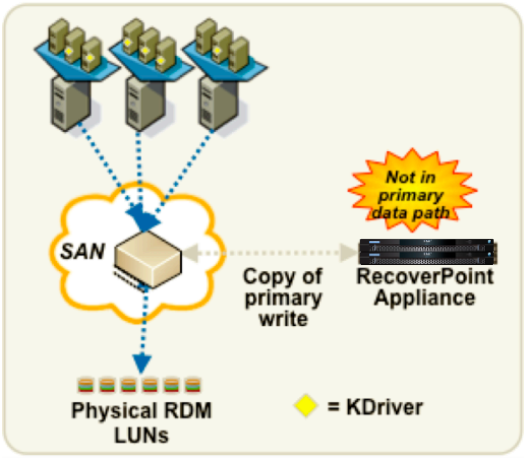

Figure 3. Host-based write splitting

Figure 3 shows that six of the nine virtual machines have a Windows-based write splitter installed. For those virtual machines, the Windows-based write splitter captures writes to each SAN volume that is attached to each guest virtual machine in physical RDM/P mode and sends a copy of those writes to the RecoverPoint appliance. These six virtual machines can have their data replicated by RecoverPoint; the other three virtual machines cannot have their data replicated by RecoverPoint.

Intelligent fabric write splitting

RecoverPoint write splitting is also provided through intelligent fabric APIs provided on EMC Connectrix switches using Brocade and Cisco technology. RecoverPoint supports the Brocade Storage Application Services APIs on the Connectrix AP-7600B switch. RecoverPoint also supports the Cisco SANTap APIs provided on the Connectrix Storage Services Module and the MDS 18/4 Multi-Services Blade, either of which can be installed in a Connectrix MDS-9000 director family, or in the Connectrix MDS-9222i switch. For replication, VMFS 3 and VMFS 5 volumes as well as volumes attached to the virtual machine in RDM/P are supported.

When a write is issued by a virtual machine it is actually handled by the hosting ESX server, which will send the write to an RDM/P volume or to a VMFS volume. The intelligent switch will identify the write target as one of the volumes that RecoverPoint is replicating and will send a copy of the write to the RecoverPoint appliance. This is an out-of-band, split-path implementation that ensures the original write is sent on to

the target with no performance impact and that reads are processed directly without flowing through the RecoverPoint appliance.

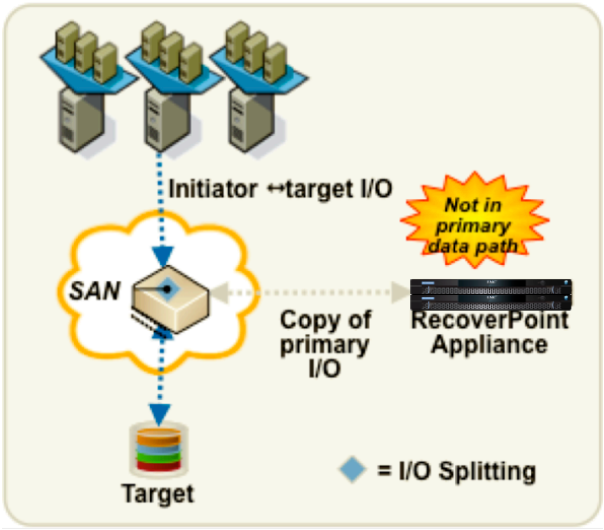

Figure 4. Intelligent fabric-based write splitting

VMware Raw Device Mapping

All of these scenarios require that VMware ESX Server storage reside on a SAN. When using a RecoverPoint host-splitter driver for a virtual machine the applications access their data using a VMware volume RDM/P in physical compatibility mode. Additionally, when using RecoverPoint to replicate data between a physical server and a virtualized environment, the virtualized servers must access the replicated data as an RDM/P volume or they must convert the volume into a VMFS image.

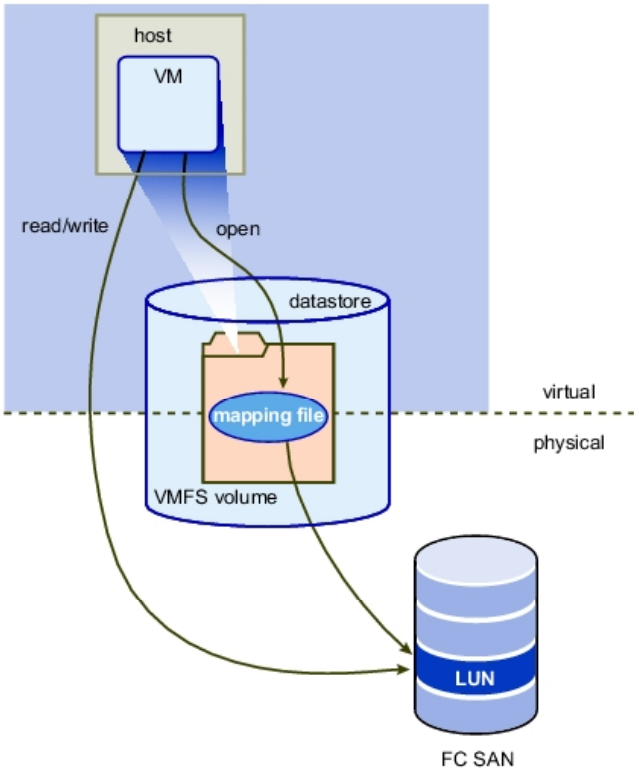

Figure 5. VMware Raw Device Mapping

Introduced with VMware ESX Server 2.5, Raw Device Mapping allows a special file in the VMFS volumes to act as a proxy for a raw device. An RDM/P can be thought of as a symbolic link from a VMFS volume to a raw LUN. The mapping makes LUNs appear as files in a VMFS volume. The mapping file, not the raw LUN, is referenced in the virtual machine configuration. When a LUN is opened for access, the mapping file is read to obtain the reference to the raw LUN. Thereafter, reads and writes go directly to the raw LUN rather than going through the mapping file.

Using RDM/P in physical mode allows SAN-aware layered applications, such as a RecoverPoint host splitter, to run inside the virtual machine. If RDM/P is not feasible, then a SAN-agnostic solution, such as the CLARiiON array-based splitter or a SAN-aware solution, such as intelligent-fabric splitters from Brocade or Cisco, must be used.

How to choose the appropriate RecoverPoint splitter

Table 1 summarizes the VMware features and limits for each of the three write-splitting technologies supported by RecoverPoint.

The simplest configuration is the host-splitter configuration. For this configuration, the RecoverPoint host splitter is installed on each virtual machine that has data that needs to be replicated. There are a few limitations on the host driver. First, it only supports the 32-bit and 64-bit Windows platforms; second, only the virtual machines' data can be replicated, and it must be attached as a RDM/P volume, and the boot

volumes are not replicated; and third, only a maximum of 255 guest machines per ESX server can be supported for replication, which is a VMware restriction.

Table 1. RecoverPoint splitter comparisons

Features / Splitter	Windows host write splitter	Array-based write splitter	Brocade/Cisco Intelligent Fabric write splitter
Supports physical RDM/P	YES	YES	YES
Supports virtual RDM	NO	YES	YES
Supports VMFS	NO	YES	YES
Supports VMotion®	NO	YES	YES
Supports HA/DRS	NO	YES	YES
Supports vSphere® 5	YES	YES	YES
Supports vCenter™ Site Recovery Manager	NO	YES	YES
Supports P2V replication	RDM/P only	RDM/P and VMFS 3 and 5	RDM/P and VMFS 3 and 5
Supports V2V replication	RDM/P only	RDM/P and VMFS 3 and 5	RDM/P and VMFS 3 and 5
Supports guest OS BFS	RDM/P only	RDM/P and VMFS 3 and 5	RDM/P and VMFS 3 and 5
Supports ESX BFS	NO	YES	YES
Maximum number of LUNs supported per ESX server	255 (VMware restriction)	N/A	N/A
Heterogeneous array support	EMC VNX, CLARiiON CX™, Symmetrix® and selected 3rd-party storage	EMC Symmetrix VMAXe, VNX and CLARiiON CX3/CX4	EMC+3rd Party
Can be shared between RecoverPoint clusters	NO	YES	NO

As this table shows, the array-based write splitter is the most effective configuration for VMware replication. With an array-based write splitter, any of the RDMs or volumes containing VMFS can be replicated. The only restriction on the write splitter is that *all* of the volumes for a specific VM must reside on arrays that are supported by and attached to the RecoverPoint appliance.

The final configuration uses intelligent fabric splitting, which is provided on some Brocade and Cisco director-class switches. Using intelligent fabric splitting the volumes can reside anywhere in the SAN. Intelligent fabric splitting enables the replication of ESX boot volumes, virtual machine boot volumes (either as physical RDM/P volumes or as a VMFS volume), and virtual machine data volumes (either as physical RDM/P volumes or as VMFS volumes). Additionally, intelligent fabric splitting is the only way that replication across heterogeneous storage arrays (EMC to third party) is supported.

VMware ESX Server

VMware ESX Server is virtual infrastructure software for consolidating and managing systems in mission-critical environments. VMware ESX Server speeds service deployments and adds management flexibility by partitioning x86 servers into a pool of secure, portable, and hardware-independent virtual machines.

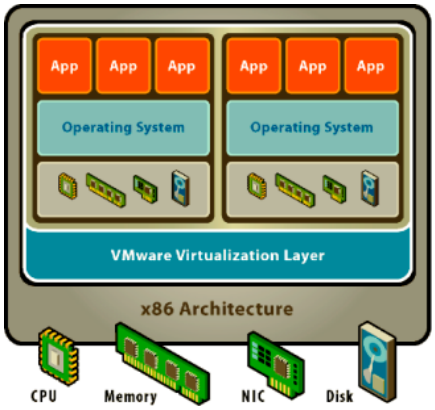

Figure 6. VMware ESX Server architecture

VMware ESX Server unifies the disaster recovery (DR) platform in a way that allows many production servers to be recovered on a single DR server without the need for costly one-to-one mapping of production and DR servers. Hardware-independent VMware ESX Server virtual machines eliminate the need to maintain identical hardware at production and DR sites.

VMware ESX Server can host multiple differing operating systems and applications that run concurrently in isolated virtual machines. System resources are dynamically allocated to each virtual machine based on need and configured service-level guarantees, providing mainframe-class control and capacity utilization of x86 servers. EMC RecoverPoint supports protection of the VMware ESX Server volumes, as well as the individual virtual machines and their data.

RecoverPoint replication for VMware

Some businesses have a challenge in maintaining high availability of their virtualized servers. In the event of a failure, the business may need to re-install the host operating system and recover the data. This works if your recovery point objective (RPO) is measured in multiple hours or days, but if you have a smaller RPO you need a more realtime solution.

RecoverPoint provides a realtime replication solution for VMware ESX and virtual machines. For asynchronous replication it uses a unique small-aperture data capture technology. For synchronous replication every change is sent to the remote site. This ensures that the VMware ESX platform is protected from data corruption and guarantees recoverability with minimum to no data loss with point-in-time crash-consistent images. For local replication in the same SAN, and synchronous remote replication between two separate SANs, every write is captured and stored in the RecoverPoint journal. For remote asynchronous replication between two separate SANs, the user specifies policies for the replication that controls the aperture size to ensure that specific RPOs are met for each consistency group. Additionally the user can establish policies that enforce synchronous remote replication or utilize the dynamic synchronous capability of RecoverPoint, which automates changes in replication state from synchronous to asynchronous.

RecoverPoint supports the replication of iSCSI and SAN-attached volumes. RecoverPoint supports RDM/P and VMFS volumes and if the VMware ESX server is configured for boot from SAN (BFS), then the boot volumes can also be replicated to the remote site. RecoverPoint captures changes to data by intercepting every write (either to an RDM/P volume or to a VMFS volume) that reaches the SAN though the intelligent fabric or CLARiiON-based splitters. Only the CLARiiON-based splitter supports iSCSI volumes. If intelligent fabric or a CLARiiON-based splitter is not used, then RecoverPoint captures the changes by using a Windows-based KDriver; however the KDriver can only capture changes written to an RDM/P volume.

To ensure that the images are consistent for each virtual machine, it is recommended that you create frequent RecoverPoint bookmarks while the ESX server is in a quiesced state. To quiesce the ESX server, first power off all guest virtual machines that reside on replication LUNs or VMFS volumes. Once the virtual machines are powered off create a RecoverPoint bookmark for the appropriate consistency groups using either the RecoverPoint GUI or CLI. Alternatively, it is possible to use the VMware Tools SYNC driver (LGTO_SYNC) to flush pending writes to a VMFS before creating the bookmark.[1] Most of today's applications and databases have a built-in resiliency allowing them to deal with crash-consistent images without the need to flush pending writes or shut down virtual machines.

RecoverPoint's image access technology allows administrators to access any image in seconds and to mount it directly as either a VMFS volume or as a RDM/P volume. Once an image is mounted it can be repurposed for backup or DR testing, or used for immediate recovery of files, folders, volumes, or entire virtual machines. When using

[1] While it is possible to use the SYNC driver, a description of this procedure is beyond the scope of this document.

RecoverPoint to replicate the data, there is no awareness of the virtual infrastructure at the destination site. You will either need to script a process to scan for and register virtual machines on the replicated volumes, or you will need to manually configure each virtual machine on the destination side. Alternatively you can utilize VMware vCenter™ Site Recovery Manager to automate this process.

Regardless of the technology, the target virtual machines will be stored in a powered off, or cold, state until they are required. To provide some level of virtual machine awareness RecoverPoint implemented a storage replication adapter, which enables RecoverPoint to be utilized as an external replication provider for VMware vCenter Site Recovery Manager. VMware vCenter Site Recovery Manager automates the scanning and registration process of the replicated volumes for the virtual machines and their data as part of the disaster recovery failover process.

VMware vCenter Site Recovery Manager

Traditional disaster recovery solutions leave many organizations unable to meet recovery time and recovery point objectives. The slow and often manual recovery processes common in traditional disaster recovery solutions are prone to errors and result in frequent failures. VMware vCenter Site Recovery Manager (SRM) provides business continuity and disaster recovery protection for virtual environments. Protection can extend from individual replicated datastores to an entire virtual site. VMware's virtualization of the data center offers advantages that you can apply to business continuity and disaster recovery.

Using VMware vCenter Site Recovery Manager with RecoverPoint

VMware vCenter SRM automates the recovery process so it becomes as simple as pressing a single button. There is no need for the user to interact with the RecoverPoint console and the VMware console; instead VMware automates the process. All the user has to do is ensure that the production virtual machines are mapped to LUNs that are replicated by RecoverPoint to the remote site.

SRM leverages an external replication solution between protected and recovery sites. The workflow that is built into SRM automatically discovers the datastores that are set up for replication between the protected and recovery sites. You can configure SRM to support bi-directional protection between the two sites.

As you can see by Figure 7, RecoverPoint sits below the VMware Infrastructure and is responsible for replicating all changes from the production LUNs to the remote replicate LUNs at the disaster recovery site. The RecoverPoint storage replication adapter is installed on the same servers that are running vCenter Server and the vCenter Server Site Recovery Manager plug-in in the production and disaster recovery sites.

The benefits of RecoverPoint are that the replication can be between differing arrays, such as between a Symmetrix and a VNX series or between an EMC storage array and a qualified third-party storage array. Additionally, there is no requirement that the production volumes be attached to the VMware servers in RDM/P mode; instead the data can reside on VMFS file systems that are contained on the production LUNs.

Finally, with RecoverPoint the distance between the sites is not a limit, since RecoverPoint replicates the data asynchronously but maintains the write-order consistency at the remote site, ensuring that all replicas remain fully consistent.

Integration with RecoverPoint

VMware vCenter Site Recovery Manager is designed as a plug-in to VMware vCenter Server so that the SRM disaster recovery tasks can be executed inside the same management tool as other VM administration tasks such as creation, migration, and deletion. VMware vCenter SRM is highly automated and is responsible for the setup, test, and recovery workflows for disaster recovery automation. SRM enables you to accelerate recovery and ensure successful recovery by automating the recovery process and eliminating the complexity of managing and testing recovery plans. VMware vCenter SRM eliminates complex manual recovery steps and removes the risk and worry from disaster recovery.

VMware vCenter SRM reduces the RTO for disaster recovery and relies on block-based replication, such as is provided by RecoverPoint, to reduce the RPO for disaster recovery. To implement replication, a RecoverPoint storage replication adapter for VMware is used to map the VMware vCenter SRM requests into the appropriate RecoverPoint actions. The RecoverPoint adapter was developed, qualified, and supported by EMC. The RecoverPoint adapter can be downloaded from the VMware site and is also available from EMC.

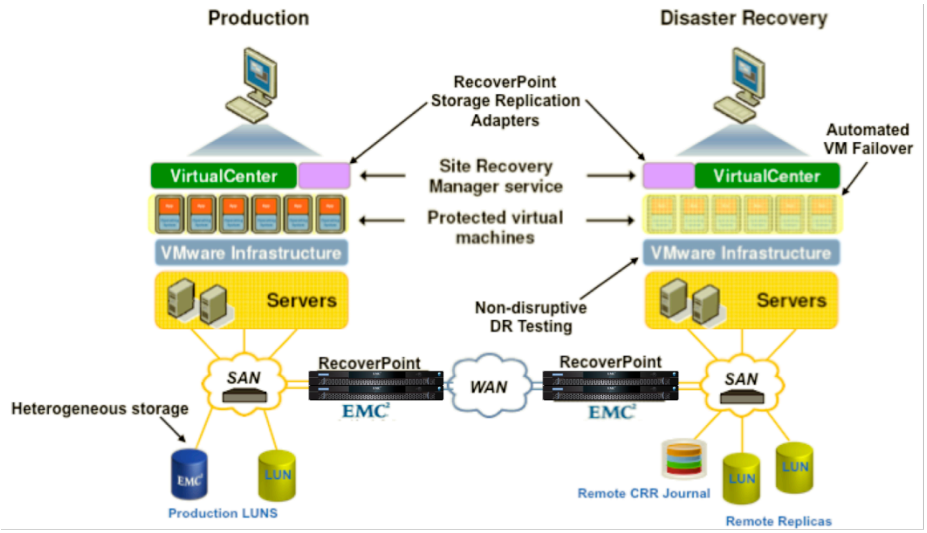

Figure 7. VMware vCenter Site Recovery Manager

All of the benefits of RecoverPoint discussed in this paper also apply when VMware vCenter SRM uses RecoverPoint for replication. This includes the use of array-based write splitters, heterogeneous storage, and policy-based replication. SRM is designed for site-to-site replication, as such it only works with remote replication as provided by RecoverPoint CRR, or if continuous local and remote data protection is being used, SRM will only operate with the remote replica copy. The local (or CDP) copy of a CLR consistency group will be unaffected by the SRM operations and will

remain available for use; however, RecoverPoint will pause the transfer to the local copy until production is resumed at the protected site. If an SRM failover has occurred and you need to access the local copy, the RecoverPoint consistency group must be changed from "Group is managed by SRM, RecoverPoint can only monitor" to "Group is in maintenance mode. It is managed by RecoverPoint." Once this is done, the local copy can be managed though the RecoverPoint management application GUI or command line. This is done automatically by products that integrate with RecoverPoint, such as EMC Replication Manager.

Ensuring reliable recovery through automation and testing

Testing disaster recovery plans and ensuring that they are executed correctly are critical to making recovery reliable. However, testing is difficult with traditional solutions due to the high cost, complexity, and disruption associated with tests. Another challenge is ensuring that staff are trained and prepared to successfully execute the complex process of recovery.

VMware vCenter SRM helps you overcome these obstacles by enabling realistic, frequent tests of recovery plans and eliminating common causes of failures during recovery. It provides built-in capabilities for executing realistic, nondisruptive tests without the cost and complexity of traditional disaster recovery testing. Because the recovery process is automated, you can also ensure that the recovery plan will be carried out correctly in both testing and failover scenarios. Site Recovery Manager also leverages VMware Infrastructure to provide hardware-independent recovery to ensure successful recovery even when recovery hardware is not identical to production hardware.

SRM is not a replication technology but a technology that manages the processes and automation steps in recovery. As such, it is dependent on other technologies, such as RecoverPoint, to replicate data from the primary site to the secondary site. SRM recovery plans can leverage the RecoverPoint image access capability to nondisruptively test the failover process to ensure that the secondary image is consistent and usable. SRM relies on two independent VMware vCenter Servers to be in place at both the protected (primary) site and at the recovery (secondary) site to facilitate the failover process between the two sites.

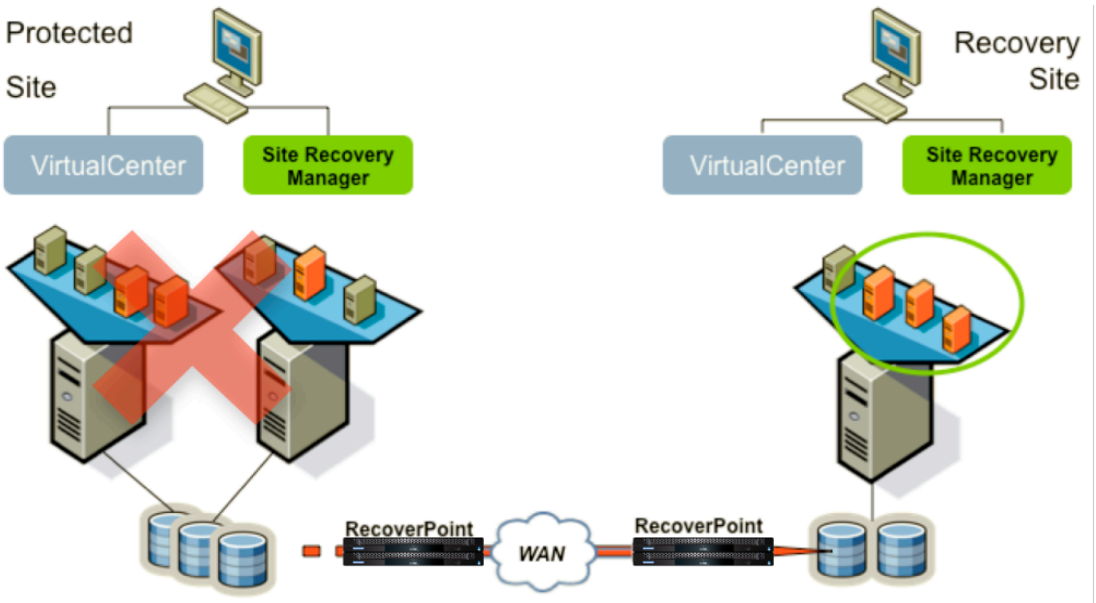

Figure 8. VMware vCenter Site Recovery Manager at a glance

The previous figure shows what the architecture looks like once SRM has been installed and configured to support the VMware Infrastructure environment. For a brief video chalk talk see *EMC and VMware: The Ultimate Disaster Recovery Solution* at http://www.emc.com/collateral/demos/microsites/mediaplayer-video/video-baker-tothepoint.htm.

In order for SRM to make use of a replication technology, a storage replication adapter must be written for that technology. An adapter was written by EMC and certified and distributed by VMware that integrates EMC RecoverPoint with VMware vCenter Site Recovery Manager. The EMC RecoverPoint adapter for VMware Site Recovery Manager is a software package that allows VMware Site Recovery Manager to implement disaster recovery using EMC RecoverPoint. The RecoverPoint adapter supports SRM functions, such as failover and failover testing, using RecoverPoint as the replication engine.

The RecoverPoint adapter supports discovery of arrays attached to RecoverPoint and discovery of RecoverPoint consistency groups that can be managed by VMware vCenter SRM, and supports SRM functions such as failover and failover testing, using RecoverPoint as the replication engine.

Site Recovery Manager failover

VMware vCenter Site Recovery Manager has rules to define which virtual machines are important and how they should be started at the protection site in the event of a disaster. SRM accomplishes this by defining the following two entities:

- **Protection Groups** are located at the protected site and define which virtual machines to protect.
- **Recovery Plans** are located at the recovery site and define steps for recovering virtual machines.

A RecoverPoint consistency group is a data set of SAN-attached storage volumes at the production site and DR site. An SRM protection group is a group of virtual machines that are failed over together (during testing and actual failover). When SRM performs failover, it instructs RecoverPoint to operate on all the LUNs of all the VMs in the protection group. RecoverPoint, on the other hand, uses consistency groups to define groups of LUNs that are replicated together. Therefore, it is recommended to:

- Group *all* SRM protection group LUNs into a single or a small number of RecoverPoint consistency groups. This grouping is to prevent SRM protection group LUNs from being left out of RecoverPoint consistency groups. If RecoverPoint is monitoring the Virtual Center at the SRM protection site, then any virtual machines that have LUNs that are not part of a RecoverPoint consistency group will be shown as partially or unprotected.
- Ensure RecoverPoint consistency groups containing SRM protection group LUNs do *not* contain non-SRM protection group LUNs. Otherwise, if such LUNs are added to the RecoverPoint consistency group, they will be handled by SRM.
- Ensure RecoverPoint consistency groups containing SRM protection group LUNs do *not* contain LUNs from more than one SRM protection group. Otherwise, SRM may attempt to operate on the same consistency group concurrently. This may cause SRM operations to fail, or to behave erratically. When these guidelines are followed then in the event of a failure at the protected site, SRM will fail over production to the protected site. As part of this event, SRM:
  - Automates failover of the EMC RecoverPoint consistency group(s)
  - Has RecoverPoint roll the physical volume(s) to the last image replicated from the production site
  - Has RecoverPoint resume production for the consistency group(s) at the temporary production site. This ensures that any changes to the production volumes at the temporary production site are replicated back to the original production site
  - Rescans and adds the datastores that reside on the volumes contained within the RecoverPoint consistency groups
  - Registers and powers on the virtual machines at the protected site

These virtual machines become the temporary production data center as shown in Figure 9.

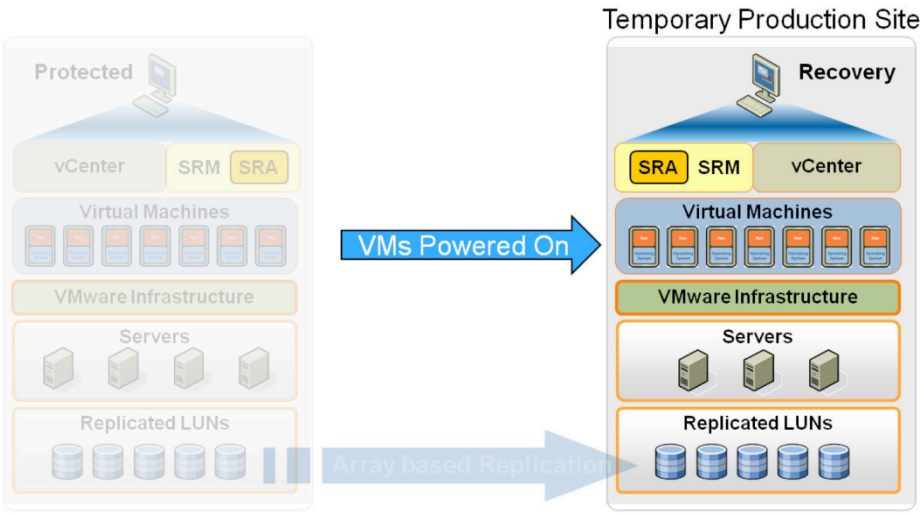

Figure 9. Site Recovery Manager failover

Using VMware SRM with RecoverPoint allows the user to automate all steps to restore production after the failover, which helps take error-prone manual steps out of the equation and makes for a smooth transition to the temporary production site.

Taking control of disaster recovery plans

Until now, keeping recovery plans and their associated processes accurate and up to date have been practically impossible due to the complexity of plans and the dynamic environment in today's data centers. Adding to that challenge, traditional solutions do not offer a central point of management for recovery plans and make it difficult to integrate the different tools and components of disaster recovery solutions.

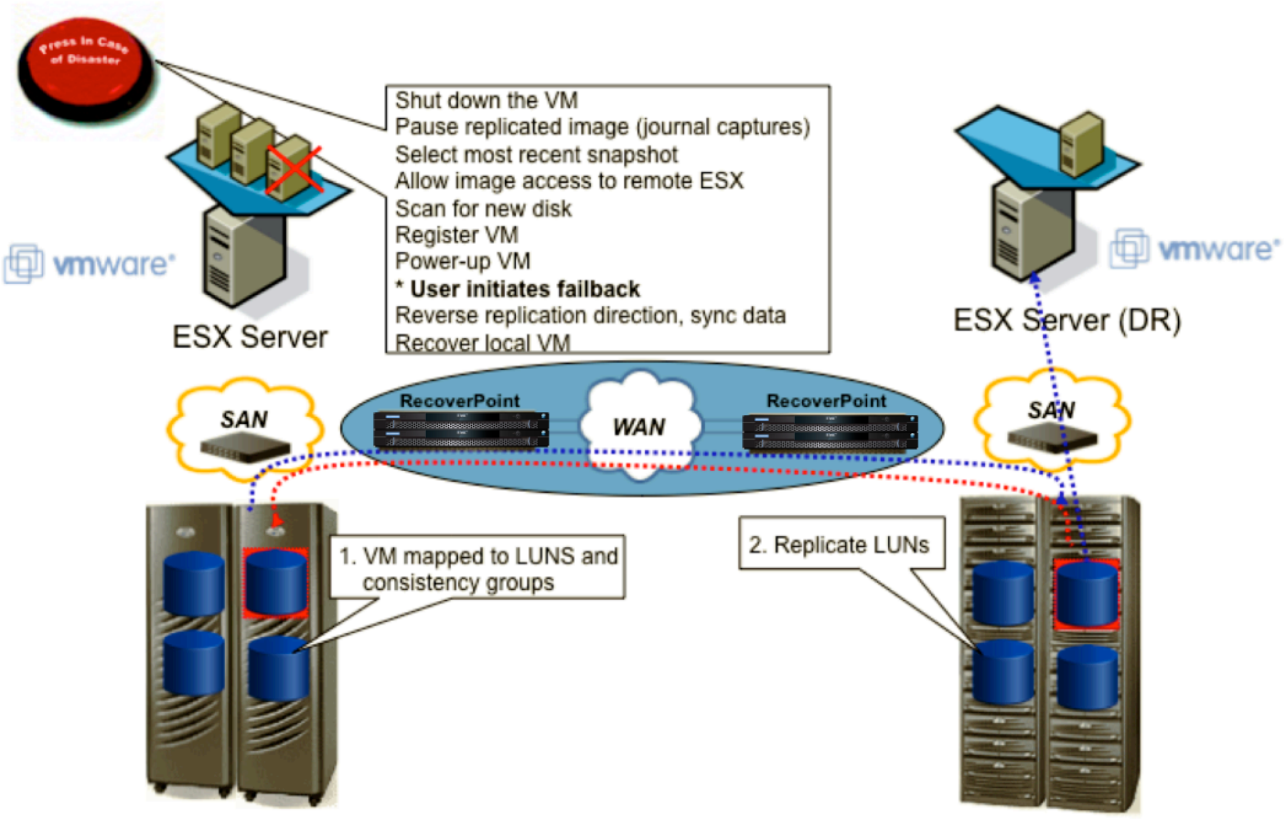

Figure 10. Automated recovery using Site Recovery Manager and RecoverPoint

VMware vCenter SRM simplifies and centralizes the creation and ongoing management of disaster recovery plans. SRM turns traditional oversized disaster recovery processes into automated plans that are easy to manage, store, and document. Additionally, SRM is tightly integrated with VMware Infrastructure 3, so you can create, manage, and update recovery plans from the same place that you manage your virtual infrastructure.

Figure 10 shows that during a disaster the VMware administrator can utilize SRM to perform the disaster recovery. Following the recovery plan, SRM shuts down the appropriate virtual machine(s) and selects the most recent image replicated by RecoverPoint. At the remote site this image becomes visible to the DR ESX server. At the point that the DR ESX discovers the new LUNs, SRM registers the virtual machines that reside on the LUNs and then powers up the virtual machines. When the administrator needs to fail back the virtual machine it is very simple to use the RecoverPoint GUI to reverse the replication direction and resynchronize the data. A semi-automated overview using Site Recovery Manager is described below. Once data is fully synchronized a failback can be performed where the virtual machine at the DR site is shut down and the local virtual machine is restarted with the recovered data.

Site Recovery Manager failback

A failback operation switches data processing from the recovery site (target) to the protected site (source). VMware vCenter SRM versions earlier than 5.0 do not support automated failback. For these releases you can perform a partially automated failback with Site Recovery Manager; a manual failback; utilize the manual failback capability written by EMC and provided in the RecoverPoint storage replication adapter; or utilize the automated site recovery manager written by EMC and provided as a vCenter plug-in. For SRM versions earlier than 5.0 VMware recommends performing failback with SRM; refer to the relevant SRM documentation. A high-level overview of the manual failback process is presented here:

1. Delete protection groups at the protection site and a recovery plan at the recovery site
2. Create protection groups at the recovery site and a recovery plan at the protection site
3. Execute a recovery plan at the protection site in test mode.
4. If the test is successful, execute a recovery plan in recovery mode.
5. Delete protection groups at the recovery site and a recovery plan at the protection site.
6. Create protection groups at the protection site and a recovery plan at the recovery site.

Note: Protection site = production/primary site; recovery site = remote/secondary site; protection group = replication group on the protection site/shadow group on the recovery site

---

The failback capabilities in the RecoverPoint SRA help automate the previous steps; additionally the capabilities of EMC RecoverPoint help simplify the recovery of production to the original protected site, by supporting bi-directional replication and automating the failback process (outside of SRM) though RecoverPoint wizards. For information about performing a failback using RecoverPoint (called Resume Protection in the RecoverPoint system) refer to the *EMC RecoverPoint Administrator's Guide.*

Automated failback for Site Recovery Manager versions earlier than 5.0

The user can use the EMC RecoverPoint Site Recovery Manager failback tool to automate the failback. This tool is a plug-in to VMware vCenter Server Virtual Center, just like VMware vCenter Site Recovery Manager. Once installed, users have access to the features and steps necessary to automate the failback process after VMware vCenter SRM has been used to mitigate a disaster. Similar to SRM, the tool helps users configure the relationship between the disaster recovery site and the production site. The user identifies the LUNs that have been failed over, and simply presses the failback button – RecoverPoint handles the rest.

This tool is not a replacement for VMware vCenter SRM and should be used with SRM. This tool does not provide users with a mechanism to execute custom script or business logic that is specific to their organization. This tool does not provide an ordering mechanism for virtual machine dependencies. It does not reset the IP addresses on each virtual machine and does not support a complete data center loss. This tool requires access to the original RecoverPoint configuration to resolve replication configuration and settings to automate failback.

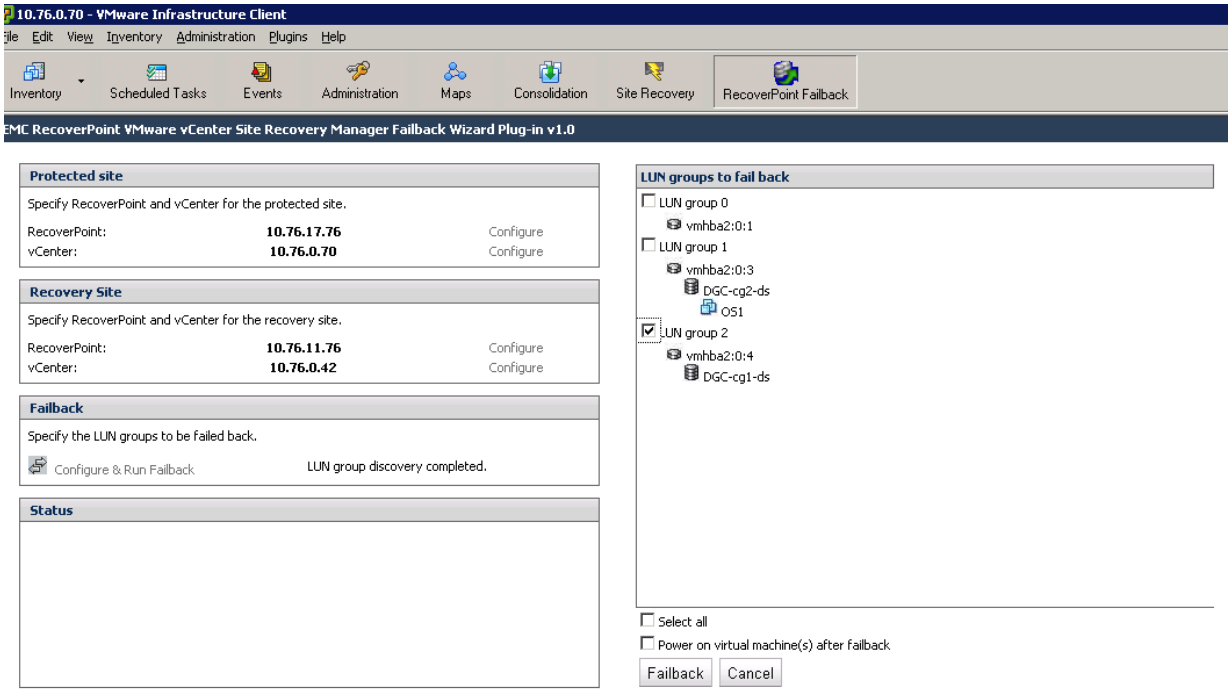

Figure 11. RecoverPoint vCenter Site Recovery Manager failback tool

Automated failback for Site Recovery Manager version 5.0

The RecoverPoint Storage Replication Adapter 2.0 with Site Recovery Manager 5.0 supports the reprotect and failback operations in VMware vCenter Site Recovery Manger 5.0. Customers that have Site Recovery Manager 5.0 must use the RecoverPoint Storage Replication Adapter 2.0 for failback operations. The EMC RecoverPoint Site Recovery Manager failback tool is not supported for these configurations.

VAAI support

VMware vSphere™ 4.1 introduces the vStorage API for Array Integration (VAAI), which speeds up certain VMware operations by offloading them to array-based hardware. By default, all VAAI commands are enabled when upgrading to or installing ESX/ESXi™ 4.1. If the RecoverPoint splitter does not support a particular VAAI command, this VAAI command must be disabled on all ESX servers in the vSphere cluster before presenting datastores to ESX hosts. The VAAI commands are:

- **Full copy:** Can significantly speed up the process of deploying virtual machines. Implemented via the **xcopy** SCSI command.
- **Block zero:** May speed up bulk zeroing of a disk. Also called **copy same.**
- **Hardware-assisted locking:** Implements a LUN locking mechanism that is more efficient in the clustered host environment. Also called **Atomic Test and Set.** Implemented using the **Compare and Swap** SCSI command.

> *Notice: Failure to disable an unsupported VAAI command may lead to data corruption, production data being unavailable to ESX hosts, degraded performance, or switch reboot.*

RecoverPoint splitters support each VAAI command at one of the following levels:

- **Unsupported.** If the RecoverPoint splitter does not support a particular VAAI command, it must be disabled on all ESX servers. Failure to disable an unsupported VAAI command may lead to data corruption, production data being unavailable to ESX hosts, degraded performance, or switch reboot.
- **Rejected.** When RecoverPoint blocks or rejects the use of a VAAI command, VMware automatically immediately reverts to legacy behavior, with no risk to data or performance.
- **Supported.** RecoverPoint supports the VAAI command and its functionality.

Table 2. VAAI commands supported by RecoverPoint, according to write-splitter type

Feature	Windows host-based write splitter	Brocade write splitter	Cisco-SANTap write splitter*	VNX-based write splitter	Symmetrix VMAXe write splitter	CLRiiON-based write splitter
Full copy	Unsupported	Rejected	Rejected	Rejected (FLARE 30) Supported (FLARE 31 or later)	Rejected	Rejected
Block zeroing	Unsupported	Rejected	Rejected	Rejected (FLARE 30) Supported (FLARE 31 or later)	Rejected	Rejected
Hardware-assisted locking	Unsupported	Rejected	Rejected	Supported	Supported	Supported

* Requires SSI 5.0(4j) and later or SSI 4.2(3k) and later in 4.2 series, prior releases or versions are unsupported

VMware affinity

In the RecoverPoint Management Application GUI the user can view ESX servers and all their virtual machines, datastores, and RDM/P drives. This view also displays the replication status of each volume

The user will see each VMware vCenter Server monitored by RecoverPoint. Under each vCenter Server object the user will see data extracted from the VMware vCenter Server, including its site name and username, its IP address, and the RecoverPoint replication state for each of the virtual machines managed by that vCenter Server. The user can easily filter out those ESX servers, VMs, and volumes that they don't want RecoverPoint to monitor. The following information is displayed for each monitored vCenter Server:

- Each ESX server in the vCenter Server and its IP address
- Each virtual machine configured and owned by the ESX server
- The replication status of each virtual machine showing if the VM is fully configured for replication, partially configured for replication, or not configured for replication
- The primary IP address of each virtual machine
- The replication status of each LUN and raw device attached to each virtual machine showing if the volume is configured for replication or not configured for replication
- If the volume is configured for replication by RecoverPoint, then the RecoverPoint consistency group, copy type (production, local, or remote), replication set that contains the LUN, and which ESX datastore contains the LUN

In Figure 12, RecoverPoint is monitoring the virtual machines that belong to a vCenter Server at 10.10.0.101. RecoverPoint discovered two ESX servers being managed by the vCenter Server. The first ESX server is at 10.10.180.10, and the other at 10.10.180.11.

The first ESX server discovered by RecoverPoint has two virtual machines (Exchange at 10.30.10.10, and LDAP at 10.30.10.11). This screen shows that RecoverPoint has discovered that the Exchange virtual machine is partially protected, with the volume (vmhba1:1:1) in the ESX Datastore: Exchange DB Datastore not protected. The other machine is called LDAP and RecoverPoint has discovered it has a single volume (vmhba:1:2:0) that is not protected. The ESX server at 10.10.180.11 has one virtual machine called LDAP (testing) at 10.40.1010 that is fully protected by RecoverPoint. This can be seen with the virtual machine's single volume (vmhba:1:3:0) in the ESX Datastore: LDAP Datastore (Testing) being fully protected by RecoverPoint.

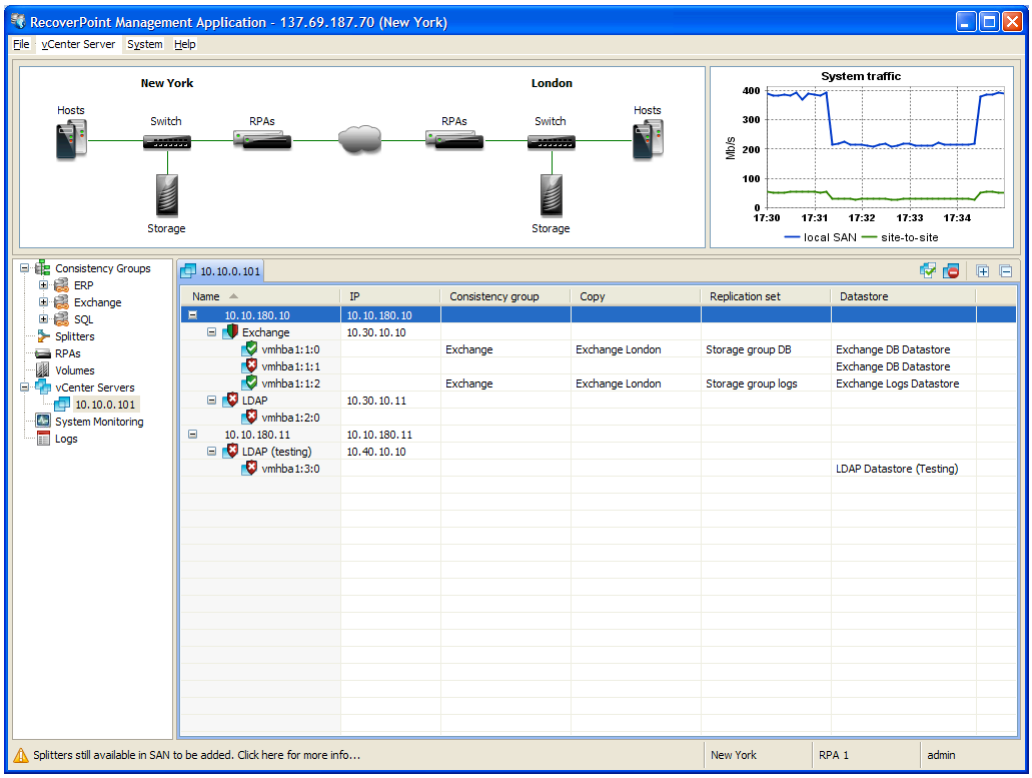

Figure 12. VMware affinity

For virtual machines that are partially protected by RecoverPoint, the administrator can modify existing RecoverPoint consistency groups, such as the Exchange consistency group, to add a replication set containing the unprotected volume. For virtual machines that are not protected by RecoverPoint, the administrator can create a new consistency group with a replication set for each volume not configured for replication.

Any change to virtual machines' protection status, such as moving from being fully protected to being partially protected, will result in RecoverPoint logging and raising an alert. Additionally, all of the information shown on the Management Application GUI for each virtual machine can be queried though the RecoverPoint CLI.

Replication configurations

The following configurations for protection of VMware ESX Server virtual machines are considered in this white paper:

- **P2V**: Replication of a physical server to a local and/or remote standby virtual machine

  In this use case, the customer has implemented VMware Infrastructure at the remote site that is utilized for disaster recovery. The production site is comprised of physical servers, or is a mix of physical servers and virtualized servers. One or more of the servers can be recovered onto a standby DR virtual machine running in an ESX server at a DR site.

- **V2V**: Replication of a virtual machine to a local and/or remote virtual machine

  In this configuration one or more virtual machines in one or more ESX servers can be recovered onto a standby ESX server at a DR site.

- **P2P with virtual CDP**: Replication of a physical server to a remote standby physical server with local replication to a standby virtual machine

  In this configuration data from local physical servers is replicated to a remote DR site. Additionally, local replication using RecoverPoint's continuous data protection technology enables the protection and importing of the data to an ESX server residing in the local SAN.

- **V2P with physical CDP**: Replication of a virtual machine in one or more ESX servers to local and remote physical machines

  In this configuration, the local virtual machine is continuously protected and its data can be recovered onto another local physical machine or on to a physical machine at the DR site. This requires that the virtual machine's data resides in an RDM/P volume attached in physical compatibility mode.

- **V2V with VMware Site Recovery Manager**: Replication of a virtual machine to a remote virtual machine with failover automation provided by VMware Site Recovery Manager

  In this configuration, one or more virtual machines in one or more ESX servers can be automatically recovered onto a standby ESX server at the DR site through the integration of RecoverPoint with VMware Site Recovery Manager.

Physical-to-virtual replication

This is a common configuration for customers that have deployed VMware Infrastructure to provide disaster recovery resources in a remote data center, but have not completed the conversion of their production data center to VMware Infrastructure. This is commonly referred to as physical-to-virtual replication. Additionally, customers that are evaluating the use of a virtualized infrastructure for their production environment can use this replication to clone their production environments and test them in a virtualized configuration.

This configuration requires that the VMware virtual machines utilize RDM/P volumes, or that VMware Converter is used to convert these volumes to VMFS.

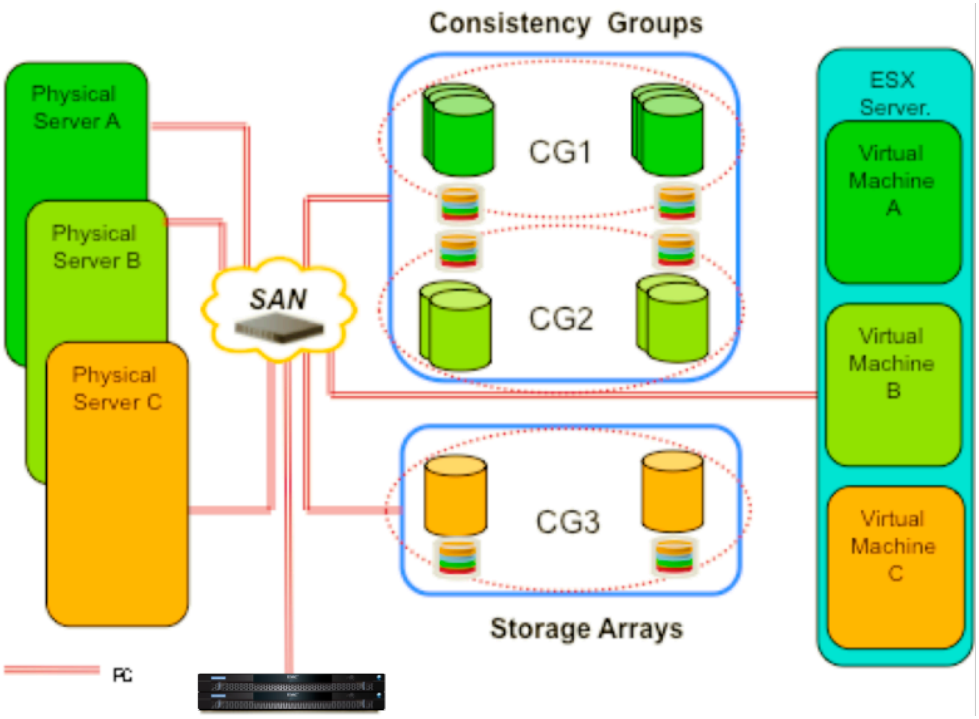

Figure 13. Physical-to-virtual replication

In physical-to-virtual replication, each physical machine is mapped to a RecoverPoint consistency group. Each of the LUNs accessed by the production machine becomes a replication set in the consistency group. Separating the physical server replication by consistency group allows for either a planned failover or for the use of the replicated image in the ESX machine for purposes such as testing, data mining, and object recovery. The target virtual machine can be either local, as shown in the previous figure, and/or remote. With RecoverPoint there is no need to have the replicas reside in the same array, array family, or volume type as RecoverPoint will manage the physical replication between dissimilar arrays.

The ESX server is built and configured with virtual machines that match the application configuration of the physical machines. The virtual machines are not running during normal operations, and are only powered on to access new data. The basic flow is as follows:

1. Select the appropriate image using either a point-in-time selection or a bookmark selection requesting physical image access. This will cause RecoverPoint to roll back the replica image to the selected point-in-time image.
2. Once the rollback is completed, the selected image LUNs will be unmasked and will become visible to the ESX server.

3. From the ESX Server console, scan for the new LUNs and then register them with the appropriate virtual machine.
4. Power up the virtual machine, which will see the point-in-time image data. At this point RecoverPoint tracks the reads and writes to the image so that they can be backed out when image access is completed.

After step 2 completes, an array snapshot, such as provided by an EMC TimeFinder® BCV or EMC SnapView™ clone, can be created from the replica LUNs, and then RecoverPoint would be informed that image access is completed. Once RecoverPoint resumes processing the user would present the snapshot to the ESX server for use by the virtual machines. Alternatively, these snapshots can be used by VMware Converter to import the data to existing VMFS volumes.

Virtual-to-virtual replication

In this usage case the customer has fully migrated to a VMware infrastructure for their production and disaster recovery sites. At the disaster recovery site, they are taking advantage of the server and storage consolidation capabilities of VMware and RecoverPoint. The customer has fully deployed a VMware ESX Server environment using VMFS deployed with either a CLARiiON-based splitter or an intelligent fabric write splitter. This configuration can also be supported using a Windows-based KDriver when SAN LUNs are attached as RDM/P volumes.

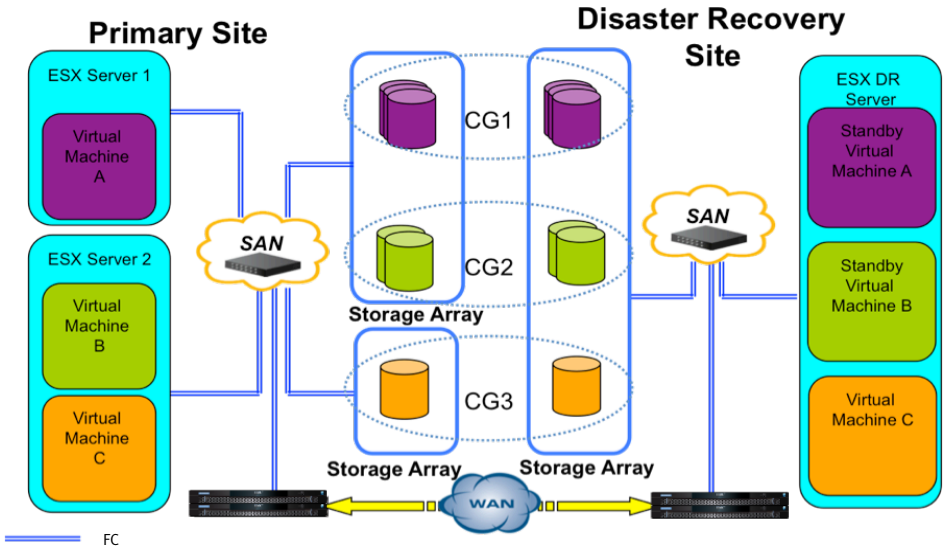

Figure 14. Virtual-to-virtual replication

If a VMFS is used, then the virtual machines are either configured to use a separate VMFS volume for each machine or all of the virtual machines use the same VMFS volumes. In the figure above, each virtual machine is assumed to have its own VMFS; this allows for an individual virtual machine to be tested and/or failed over to the disaster recovery site without impacting the remaining virtual machines. When multiple guest operating systems (OSs) are distributed over a limited set of LUNs, such as combined into one or more VMFS volumes, a special configuration is

recommended that allows single guest OS replication granularity, thereby expediting failover and test activities.

In the combined VMFS configuration, two RecoverPoint consistency groups will be defined. The first RecoverPoint consistency group contains all of the existing LUNs (or VMFS volumes) with the entire set of guest OSs and their data. The second RecoverPoint consistency group, or *standby*, contains storage volumes at each site that are at least as large as the largest combined guest OS image and related virtual data disks. Although this group is configured for RecoverPoint replication, it does not replicate any data as long as there are no writes taking place at the source site. When needed, any guest OS with its data can be moved to this storage space, and individually replicated to the target site. As a result, it is also possible to test and, if desired, fail over this guest OS independently from the other guest OSs.

Physical-to-physical replication with local virtualization

In a physical-to-physical configuration the customer is looking to migrate from a physical to virtual environment either for their production data center or for their disaster recovery data center. Either the local replica is to be presented to a virtualized environment or the remote replica is to be presented to the virtualized environment. The example below uses RecoverPoint concurrent local and remote data protection. This differs from the local physical-to-virtual replication configuration that uses RecoverPoint continuous remote replication. This example adds the physical machines that exist in the disaster recovery site. Since physical machines exist, the virtualized machines access the data in RDM/P mode.

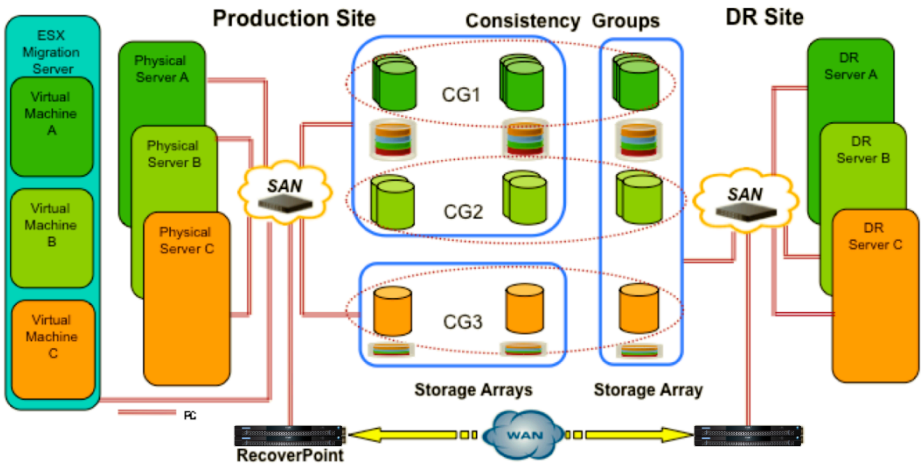

Figure 15. Physical to local and remote virtual replication

The benefit of using RecoverPoint concurrent local and remote data protection is that the replica copies can be accessed without impacting the protection and disaster recovery of the physical servers. As this configuration utilizes SAN volumes accessed as RDM/P volumes, any of the RecoverPoint write-splitting technologies can be implemented, including host, CLARiiON, and intelligent-fabric splitters.

Conclusion

The innovative technology of EMC RecoverPoint supports flexible levels of protection, without distance limitations and performance degradation. With its unique architecture, powerful data recovery features, and business-driven approach, RecoverPoint offers superior levels of local and remote data protection and business continuity to organizations running VMware ESX. Organizations implementing RecoverPoint with VMware ESX Server are expected to see the following benefits:

- Flexible storage configurations when VMware vCenter Site Recovery Manager is used with the RecoverPoint storage replication adapter
- Support for automating the SRM failback when the RecoverPoint SRA is used with SRM 4.x
- Integration with VMware vCenter Server enables the user to view and update the protection status of virtual machines
- Ability to leverage RecoverPoint concurrent local and remote data replication to accelerate the transitioning or migration to a VMware infrastructure
- Full support for VMware physical-to-virtual and virtual-to-virtual replication models
- Support for heterogeneous storage reduces the need to perform data migration between storage architectures

- Qualified with VMware technologies, including DRS, HA, Storage VMotion, and VMotion
- Support for replication between VMFS volumes as well as between RDM/P volumes
- Simple and quick planned or unplanned failover for virtual machines and their data without distance limitation
- Out-of-band processing for replication that ensures that the performance of ESX Server and its virtual machines are not impacted by RecoverPoint
- Innovative compression algorithms and intelligent bandwidth policy management that eliminate the need for dedicated IP or FC links between sites
- Rapid and simple replication for virtual machines and their data to an alternate location and instantly accessible for disaster recovery or for recovery from logical corruption
- Ability to leverage local replication for operational or application recovery of a virtual machine while still maintaining remote replication to provide protection in case of a site-wide disaster

References

More information on EMC RecoverPoint can be found at the RecoverPoint page on EMC.com and in the following documents on the EMC Powerlink website:

- Introduction to EMC RecoverPoint 3.4: New Features and Functions — Applied Technology
- Improving Microsoft Exchange Server Recovery with EMC RecoverPoint — Applied Technology
- EMC RecoverPoint Family Overview — A Detailed Review
- *EMC RecoverPoint Replicating VMware ESX Technical Notes* (Powerlink only)
- Using EMC RecoverPoint Concurrent Local and Remote for Operational and Disaster Recovery – Applied Technology
- Solving Data Protection Challenges with EMC RecoverPoint – Best Practices Planning
- *EMC RecoverPoint Adapter for VMware vCenter Site Recovery Manager Release Notes* (Powerlink only)